Adaptive data streams fuzzy clustering with an unknown number of classes*

Alina Shafronenko[1,†]

[1]Kharkiv National University of Radio Electronics, Nauky ave 14, Kharkiv, 61166, Ukraine

Abstract

This article addresses the complex problem of performing fuzzy clustering on data streams in an online, self-learning mode where the total number of classes is a priori unknown and may change over time. To address this, the study proposes a computationally efficient ensemble approach comprising parallel subsystems, each designed to identify a distinct number of fuzzy clusters. The methodology utilizes a recurrent version of credibilistic fuzzy clustering, which improves upon traditional fuzzy C-means methods by reducing sensitivity to anomalous outliers and preventing the merging of clusters into single large classes. To determine the optimal number of classes dynamically, the system evaluates the quality of results from each ensemble member using the fuzzy partition coefficient and selects the model with the maximum value. This approach is specifically designed to handle non-stationary data streams that arrive at high frequencies.

Keywords

Data Streams, adaptive fuzzy clustering, ensembles

1. Introduction

The problem of clustering-classification without a teacher is an integral part of Data Mining and, at the same time, the most challenging, as it is implemented in a self-learning mode, that is, without an external training signal [1-3].

Solving the problem becomes significantly more complicated if the data sample to be clustered is not provided in a traditional batch with a fixed number of observations, but rather is delivered online as a data stream that arrives one observation at a time, possibly at high frequency. Such a problem is addressed within the scientific direction of Data Stream Mining [4-6], which is currently undergoing intensive development.

When solving real-world, practical problems, a situation often arises in which an observation vector from a data stream can simultaneously belong to two or more classes, with varying levels of fuzzy membership. Such a situation is considered within the framework of Fuzzy Clustering Analysis [3], which also exists in the recurrent online version [7].

And finally, the most challenging situation arises when the data are presented as a stream, the classes are cross-sectional, and the total number of classes is unknown and may change over time. The most appropriate here is to use an ensemble approach, when each member of the ensemble is oriented to a different number of possible fuzzy clusters, but the existing approaches are quite cumbersome from a computational point of view, which makes their use difficult when solving real practical problems. Therefore, it is advisable to develop fast ensemble fuzzy clustering procedures

* CMIS-2026: Nineth International Workshop on Computer Modeling and Intelligent Systems, May 5, 2025, Zaporizhzhia, Ukraine

[1*] Corresponding author.
[†] These authors contributed equally.
alina.shafronenko@nure.ua (A. Shafronenko)
0000-0002-8040-0279 (A. Shafronenko)

that enable solving the problem of online fuzzy clustering of data streams under a priori uncertainty about the possible number of classes in the available observations.

2. Neuro-fuzzy system ensembles for solving clustering problems with an unknown number of classes

In the simplest case, clustering a multidimensional data stream $x(1), x(2), \ldots, x(k), \ldots x(N)$, (here $x(k) = (x_1(k), \ldots x_i(k), \ldots, x_n(k))^T \in R^n, k = 1,2, \ldots, N, N+1$- current discrete time index) a priori unknown number of classes-clusters $m = 1,2, \ldots, M$ can be implemented using an ensemble of Kohonen's self-organizing maps (SOM) [9], each of which contains the appropriate number of neurons, i.e. the first SOM is oriented to work under the assumption of two classes, the second – three, and finally the last SOM of the ensemble contains M neurons, i.e. it splits the data stream into M classes. In this case, either a set of observations on a sliding window or a vector of autocorrelations of the analyzed stream, which are also calculated online, is used as the input signal.

Let's consider the process of self-learning m-th SOM, that contains m-neurons, each of which defines a prototype centroid j-th ($j = 1,2, \ldots, m$) clusters $c_j^m = (c_{j1}^m, c_{j2}^m, \ldots, c_{jn}^m)^T$, which is tuned using standard WTA or WTM – Kohonen self-learning rules. When the input signal is fed to the m-th SOM input $x(k)$ for each neuron firstly the Euclidean distance is calculated

$$D_j^m(x(k), c_j^m(k-1)) = \|x(k) - c_j^m(k-1)\|$$

(here $c_j^m(k-1)$ - vector of synaptic centroid weights) j-th neuron m-th SOM ensemble calculated on the data $(k-1)$ previous observations, assuming that all input data are previously normalized to n-dimension hypersphere with a unit radius, on the basis of which the "winning" neuron is located $c_j^{m*}(k-1)$ nearest in mean D_j^m to incoming observation $x(k)$.

After that, the synaptic weights of the "winner" are tunned using WTA – the self-learning rule (Winner Takes All) in the form

$$c_j^m(k) = \begin{cases} c_j^m(k-1) + \eta(k)\left(x(k) - c_j^m(k-1)\right), \; if \;\; c_j^m(k-1) = c_j^{m*}(k-1), \\ c_j^m(k-1) \;\; otherwise \end{cases} \tag{1}$$

here $0 < \eta < 1$.

It is easy to see that in this case the "winner" is pulled up to the input vector $x(k)$, and at $\eta(k) = \frac{1}{k}$ the vector of corresponding synaptic weights is simply the arithmetic mean of observations from the j-th class, i.e., in fact, such a self-organizing map implements the popular K-means clustering algorithm [10] in an online version, which is most convenient when working with data streams, and minimizes the self-learning criterion

$$E(x(k), c_j^m) = \sum_k \sum_{j=1}^m \|x(k) - c_j^m\|^2.$$

Compared to the WTA self-learning rule (1), the WTM (Winner Takes More) rule is more effective, using the so-called neighborhood function, which specifies the "closeness" between each observation $x(k)$ and each centroid vector $c_j^m(k-1)$. Since the choice of this function is usually subjective, it is convenient to use a similarity measure in the form:

$$\text{¿}\; \left(x(k), \; c_j^m(k-1)\right) = x^T(k) c_j^m(k-1) = \cos\left(x(k), c_j^m(k-1)\right). \tag{2}$$

Using the similarity measure (2) as a neighborhood function leads to a modified WTM – self-learning rule

$$c_j^m(k)=c_j^m(k-1)+\eta(k)\cos\left(x(k),c_j^m(k-1)\right)\left(x(k)-c_j^m(k-1)\right)=c_j^m(k-1)+\eta(k)x^T(k)c_j^m(\cdot \qquad (3)$$

It is interesting to note here that in the process of self-learning using algorithm (3) to the image $x(k)$ centroids of the "winners" are pulled up $c_j^m(k-1)$ sufficiently close to it, that is, those for which

$$x^T(k)c_j^m(k-1)>0.$$

Neurons - "loosers" for whom

$$x^T(k)c_j^m(k-1)<0$$

repelled from the vector $x(k)$, increasing the distance from it.

The use of self-learning clustering algorithms (1), (3) assumes that each observation from the data stream $x(k)$ can belong to only one class-cluster.

A more realistic situation occurs when classes overlap, and observations can belong to multiple classes simultaneously. Within the framework of fuzzy clustering, the most popular method is the fuzzy C-means [2,3], primarily due to its simple numerical implementation. The main disadvantage of this approach is its sensitivity to anomalous outliers, which cannot be eliminated when analyzing data streams whose future behavior is difficult to predict.

More robust in this sense is the method of possibilistic fuzzy clustering [11], which, at the same time, suffers from the problem of coincidence, where, during the data stream processing, clusters begin to merge into a single large class. These shortcomings are largely eliminated by the credibilistic approach to fuzzy clustering [8, 12], based on the theory of credibility [12-14], where, in the process of processing data sets, not only the cluster centroids and the levels of fuzzy membership of each observation to each cluster are calculated, but also the credibility level in the results obtained.

In [12], recurrent fuzzy credibilistic clustering procedures were introduced, enabling the processing of data streams in an online mode. It is advisable to develop this approach for ensemble systems operating under the assumption of an unknown number of classes, or in situations where the number of clusters may change during the processing of the information stream.

Credibilistic fuzzy clustering algorithms are associated with the minimization of the criterion

$$E\left(Cr(k),c_j^m\right)=\sum_k\sum_{j=1}^{m}Cr_j^\beta\left(D_j^m(x(k),c_j^m)\right)^2=\sum_k\sum_{j=1}^{m}Cr_j^\beta\left\|x(k)-c_j^m\right\|^2$$

(here β - non-negative fuzzification parameter that specifies the blurring of cluster boundaries) in the presence of constraints

$$\begin{gathered}0\le Cr_j(k)\le 1\ \forall j,k;\\ \sup Cr_j(k)\ge 0.5\ \forall k;\\ Cr_j(k)+\sup Cr_j(k)=1\end{gathered}$$

for all j and k for whom $Cr_j(k)\ge 0$, where $Cr_j(k)$ level of confidence that the vector $x(k)$ belongs to a cluster Cl_j .

Typically, in fuzzy credibilistic clustering algorithms, the level of fuzzy membership is determined by some bell-shaped membership function [15]

$$U_j(k)=\mu_j\left(D_j(x(k),c_j)\right)$$

which monotonically decreases on the interval $(0,\infty]$, in this case $\mu_j(0)=1$ and $\mu_j(\infty)\to 0$. Thus, in [5] it was proposed to use a fairly simple construction as such a function

$$U_j(k)=\frac{1}{1+D_j^2(x(k),c_j)},$$

although in principle it is possible to use other kernel functions.

In its final form, the batch algorithm for fuzzy credibilistic clustering can be written as follows:

$$\begin{cases} U_j^m(k)=\left(1+D_j^2\left(x(k),c_j^m(k-1)\right)\right)^{-1}, \\ U_j^{m*}(k)=U_j^m(k)\left(\sup U_l^m(k)\right)^{-1}, l\neq j, \\ Cr_j(k)=0.5\left(U_j^{m*}(k)+1-\sup U_l^m(k)\right), \\ c_j^m(k)=\dfrac{\sum_{k=1}^{N} Cr_j^{\beta}(k)x(k)}{\sum_{k=1}^{N} Cr_j^{\beta}(k)}. \end{cases} \tag{4}$$

Procedure (4) implements batch processing of information, when the entire sample is given a priori and does not change during its processing.

In the case when information is received online in the form of a stream, a recurrent version of (4) can be used in the form

$$\begin{cases} \delta_j^m(k)=\sum_{\substack{l=1\\ l\neq j}}^{m}\left(D_j^2\left(x(k),c_j^m(k-1)\right)^{\frac{1}{1-\beta}}\right)^{-1}, \\ U_j^m(k)=\left(1+\dfrac{\left(D_j^2\left(x(k),c_j^m(k-1)\right)\right)^{\beta-1}}{\delta_j^m(k)}\right)^{-1}, \\ U_j^{m*}(k)=\dfrac{U_j^m(k)}{\sup U_j^m(k)}, l\neq j, \\ Cr_j(k)=0.5\left(U_j^{m*}(k)+1-\sup U_l^m(k)\right), \\ c_j^m(k)=c_j^m(k-1)+\eta(k)Cr_j^{\beta}(k)\left(x(k)-c_j^m(k-1)\right) \end{cases} \tag{5}$$

where $U_j^{m*}(k)$ - normalized values of the corresponding membership levels.

In the most common case, when fuzzification takes the value $\beta=2$ (J. Bezdek's fuzzy C-means method [16]) (5) takes a simplified and convenient form:

$$\begin{cases} \delta_j^m(k)=\sum_{\substack{l=1\\ l\neq j}}^{m}\left(\left\|x(k)-c_j^m(k-1)\right\|^2\right)^{-1}, \\ U_j^m(k)=\left(1+\dfrac{\left\|x(k)-c_j^m(k-1)\right\|^2}{\delta_j^m(k)}\right)^{-1}, \\ U_j^{m*}(k)=\dfrac{U_j^m(k)}{\sup U_j^m(k)}, l\neq j, \\ Cr_j(k)=0.5\left(U_j^{m*}(k)+1-\sup U_l^m(k)\right), \\ c_j^m(k)=c_j^m(k-1)+\eta(k)Cr_j^{\beta}(k)\left(x(k)-c_j^m(k-1)\right). \end{cases} \tag{6}$$

It is easy to see that the credibilistic clustering algorithm (6) is not computationally more complex than other similar procedures.

Thus, the ensemble consists of M-1 similar blocks, each of which solves the fuzzy clustering problem for a different number of possible classes. In order to obtain the final result, that is, to determine how many classes the analyzed sample contains, it is advisable to use one or another clustering quality index [18].

Among such indices, one of the simplest and most effective is the so-called partition coefficient, which is easy to calculate in the form

$$PC_m(k)=\frac{1}{2}\sum_{\tau=1}^{k}\sum_{j=1}^{m}U_j^m(\tau).$$

This coefficient has a transparent physical meaning: the better formed the clusters, the greater is the value $PC_m(k)$.

Minimum value of $PC_m=m^{-1}$ is achieved if the data belong equally to all clusters, i.e. clusters are not formed at all. The advantage of using this coefficient is that it is well suited for online calculation, when data can arrive with high frequency.

As a result of the ensemble's work, a member is selected for whom $PC_m(k)$ acquires its maximum value. At the same time, if the sample-data stream is non-stationary and the number of clusters can change over time, it is the ensemble approach that allows us to detect such changes.

3. Experimental studies

To evaluate the performance efficiency of the developed method and demonstrate its benefits over the analogs, an experimental study was conducted using a medical dataset. For analysis, information on diagnostic signs of cancer diseases was taken: texture, area, smoothness, compactness, prominence, radius, symmetry, fractal dimension, number of areas, etc.

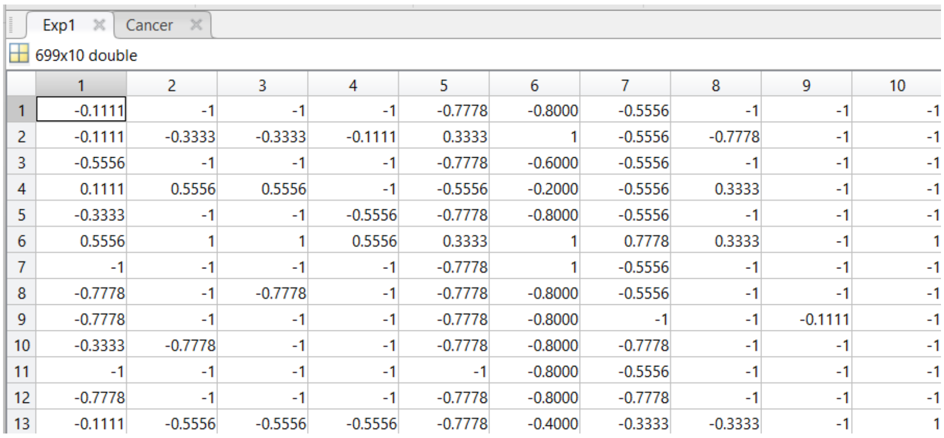
The information is presented in the form of an "object-property" table (Fig. 1). For further data mining, the sample was normalized into a hypercube [-1;1] (Fig. 2).

Exp1 Cancer

699x10 double

	1	2	3	4	5	6	7	8	9	10
1	5	1	1	1	2	1	3	1	1	2
2	5	4	4	5	7	10	3	2	1	2
3	3	1	1	1	2	2	3	1	1	2
4	6	8	8	1	3	4	3	7	1	2
5	4	1	1	3	2	1	3	1	1	2
6	8	10	10	8	7	10	9	7	1	4
7	1	1	1	1	2	10	3	1	1	2
8	2	1	2	1	2	1	3	1	1	2
9	2	1	1	1	2	1	1	1	5	2
10	4	2	1	1	2	1	2	1	1	2

Figure 1: Diagnostic signs of cancer.



Exp1 | Cancer

699x10 double

	1	2	3	4	5	6	7	8	9	10
1	-0.1111	-1	-1	-1	-0.7778	-0.8000	-0.5556	-1	-1	-1
2	-0.1111	-0.3333	-0.3333	-0.1111	0.3333	1	-0.5556	-0.7778	-1	-1
3	-0.5556	-1	-1	-1	-0.7778	-0.6000	-0.5556	-1	-1	-1
4	0.1111	0.5556	0.5556	-1	-0.5556	-0.2000	-0.5556	0.3333	-1	-1
5	-0.3333	-1	-1	-0.5556	-0.7778	-0.8000	-0.5556	-1	-1	-1
6	0.5556	1	1	0.5556	0.3333	1	0.7778	0.3333	-1	1
7	-1	-1	-1	-1	-0.7778	1	-0.5556	-1	-1	-1
8	-0.7778	-1	-0.7778	-1	-0.7778	-0.8000	-0.5556	-1	-1	-1
9	-0.7778	-1	-1	-1	-0.7778	-0.8000	-1	-1	-0.1111	-1
10	-0.3333	-0.7778	-1	-1	-0.7778	-0.8000	-0.7778	-1	-1	-1
11	-1	-1	-1	-1	-1	-0.8000	-0.5556	-1	-1	-1
12	-0.7778	-1	-1	-1	-0.7778	-0.8000	-0.7778	-1	-1	-1
13	-0.1111	-0.5556	-0.5556	-0.5556	-0.7778	-0.4000	-0.3333	-0.3333	-1	1

Figure 2: Normalized data.

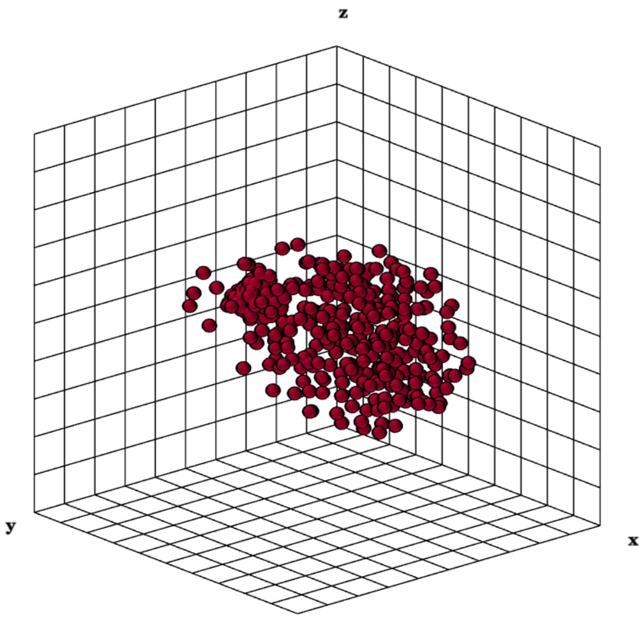


Figure 3: Scatter of observations in the space of a hypercube.

The task was previously set to divide patient observations into 2 classes: benign and malignant tumors. Fig. 4 shows a cluster analysis of the diagnostic data, where the features are divided into 2 clusters. The clustering quality is given in Table 1.

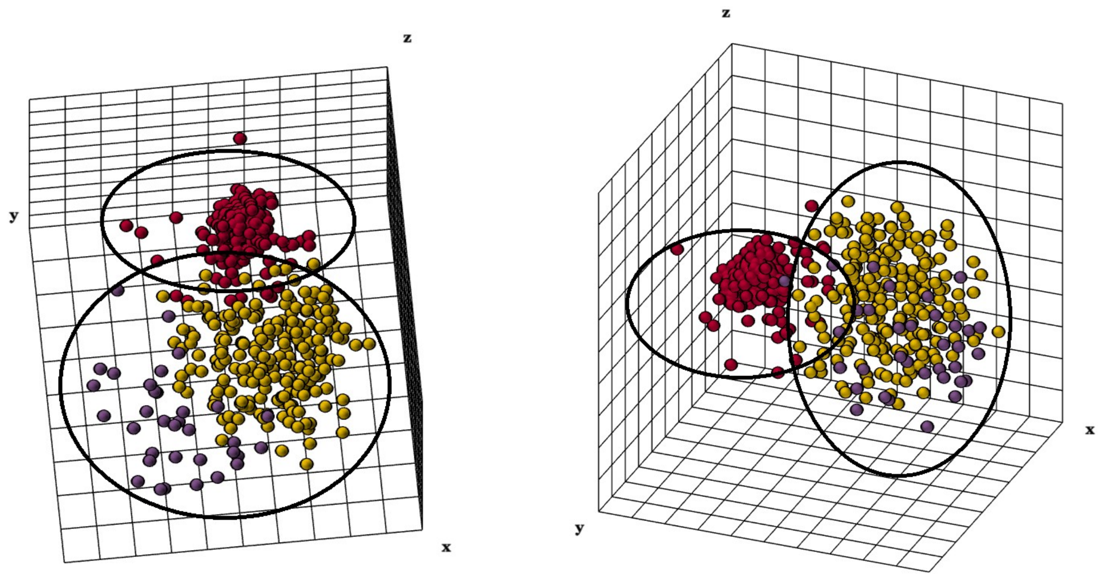


Figure 4: Splitting diagnostic data into 2 clusters (demonstration from different angles).

Table 1
Assessing the quality of data clustering

Methods	PC	CE	SC	S	XB	DI
Neuro-fuzzy system ensembles based on credibilistic fuzzy clustering algorithms	0.66	1.56	7.38	2.71	5.71	0.20
FCM	0.79	0.38	7.33	-6.84	5.61	0.01
Gustafson-Kessel	0.55	0.63	8.59	0.04	1.07	0.10

Analyzing the sample using the fuzzy data clustering method based on optimization procedures that can work even with those observations that are still completely filled (no signs, no analysis has been performed), and analyzing them.

Using the proposed method, another class-cluster was discovered. It is proposed to divide the data sample into 3 clusters, assess the clustering quality, and analyze each cluster separately. Fig. 5 demonstrates clustering of the data into 3 clusters, including observations whose information was not analyzed.

Table 2
Assessing the quality of data clustering

Methods	PC	CE	SC	S	XB	DI
Neuro-fuzzy system ensembles based on credibilistic fuzzy clustering algorithms	0.75	0.36	7.50	2.61	5.51	0.17
FCM	0.69	0.45	7.23	0.44	5.41	0.01
Gustafson-Kessel	0.45	0.53	7.24	0.04	1.05	0.10

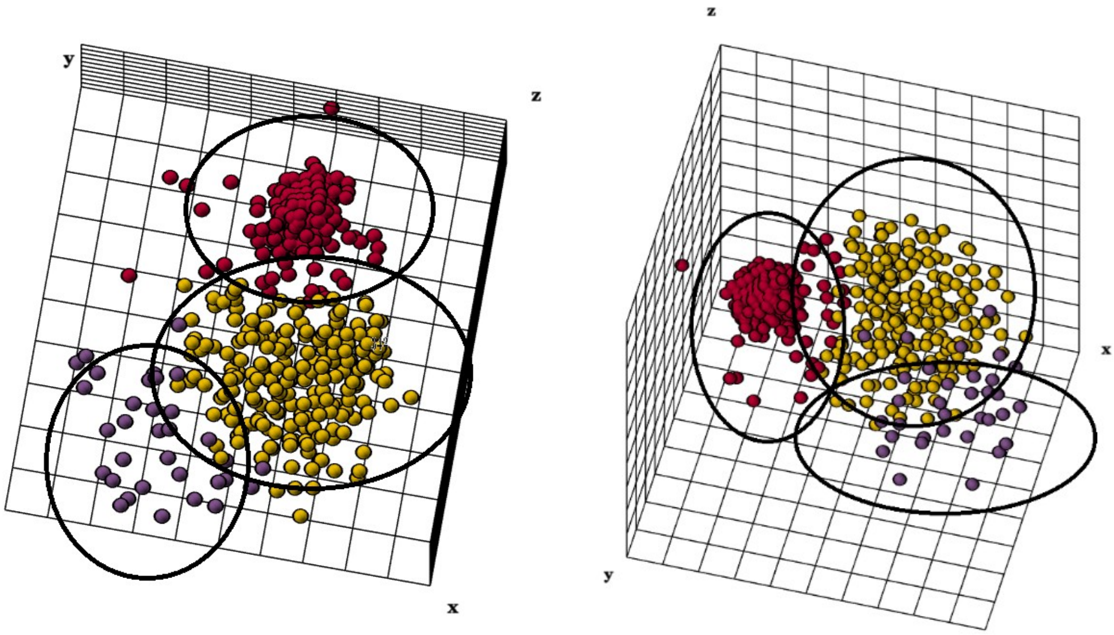


Figure 5: Dividing a sample of diagnostic data into 3 clusters (demonstration from different angles).

Analyzing Table 2, we conclude that the quality of data clustering by the proposed fuzzy clustering method yields better solutions for the relationship between a specific building and its corresponding responsibility class, which currently plays a significant role in planning work on objects. It should be noted that the main clustering quality coefficients, such as PC, SE, and XB, are more thoroughly attribute the corresponding object to the cluster class, assess the accuracy of assigning a specific observation to the cluster class, thereby removing ambiguity in the object's relationship to the cluster. Among the comparative clustering methods, none of the classical methods achieved the data clustering quality provided by the adaptive fuzzy clustering method.

Clarity and credibility of the partition: The credibilistic approach enabled us to achieve the highest separation coefficient, $PC = 0.75$, and the lowest fuzzy entropy indicator, $CE = 0.36$. This indicates high confidence in the system's ability to assign objects to specific clusters and to minimize information uncertainty, compared to the FCM and Gustafson-Kessel methods.

Cluster separability: The Dunn index score ($DI = 0.17$) for neuro-fuzzy ensembles is the highest among the studied methods, confirming their ability to better separate compact data groups in feature space.

Specificity of classical methods: It is established that the Gustafson-Kessel algorithm, despite lower overall precision ($PC = 0.45$), yields the lowest values of the deviation indices ($XB = 1.05$; $S = 0.04$). This makes it appropriate for use in problems where the priority is to account for the specific geometry and ellipsoidal shape of clusters.

To assess the effectiveness of the developed method and demonstrate its advantages over analogues, an experimental study was also conducted on a sample of water quality, its condition, and pollution.

To determine the suitability of water for agricultural use, the type and likelihood of salinity during prolonged irrigation, and to provide recommendations for improving water characteristics and/or reducing the negative impact of use, water samples from the Kharkiv region were analyzed.

Depending on the quality of the water and the required degree of treatment to meet the "Drinking Water" indicators, water bodies suitable as sources for domestic drinking water supply are divided into 3 classes according to the drinking water standards.

Despite the fact that recently there has been a trend towards a decrease in the volume of water use for the needs of sectors of the national economy, and therefore, a corresponding decrease in the volume of general drainage, the share of polluted wastewater in return waters is quite high, which ultimately causes significant pollution of water bodies with wastewater. The information was formulated as an “object-property” table, pre-normalized to a hypercube [-1;1], and cluster analysis was performed based on the three main classes (Fig. 6).

Exp1

118x38 double

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.4196	-0.8235	0.1111	-1	-0.2731	-0.8889	0.5052	-0.6525	0.3616	0.1111	-1	-0.8466	0.5887	-0.7879	0.44
2	0.0223	-0.6471	-0.1111	-1	-0.1218	-0.7822	0.6055	-0.4831	1	-0.3333	-1	-0.8211	0.7730	-0.6545	
3	-0.5093	-0.4118	-0.3333	-1	0.2353	-0.8444	0.6298	-0.7458	-0.1538	-0.3333	-1	-0.8594	0.6773	-0.8485	-0.20
4	-0.2907	-0.5882	0.3333	-0.0487	0.4874	-0.8311	0.4810	-0.6525	0.7330	-0.1111	-0.0387	-0.7827	0.6915	-0.6061	0.64
5	-0.1420	-0.8235	0.5556	0.1230	0.1008	-0.8667	0.4533	-0.6949	0.3616	0.1111	-1	-0.8339	0.1418	-0.8485	0.32
6	-0.0130	-0.6471	0.1111	-0.0626	-0.4202	-0.8444	0.5917	-0.6525	-0.1793	-0.1111	-1	-0.8147	0.4433	-0.6061	0.00
7	0.1859	-0.2941	0.1111	-1	0.3361	-0.6756	0.4291	-0.6102	-0.2269	-0.1111	-1	-0.6997	0.4929	-0.7273	-0.18
8	-0.2059	-0.4118	-0.1111	-0.0023	0.0714	-0.5156	-0.3806	-0.5254	-0.2153	-0.5556	-1	-0.7252	0.0106	-0.6364	-0.24
9	-0.7497	-0.7059	-0.1111	-0.0441	-0.0882	-0.8267	0.6021	-0.6525	-0.6378	-0.3333	-0.1609	-0.8594	0.2908	-0.8788	-0.66
10	0.0397	-0.7647	0.1111	-0.2019	0.1429	-0.8133	0.5986	-0.6102	0.0772	-0.1111	-0.1527	-0.8147	0.4433	-0.7273	0.21
11	0.2859	-0.9176	0.3333	-0.1230	0.0252	-0.7467	0.5294	-0.5678	-0.4510	0.5556	-0.2953	-0.9042	0.3156	-0.8182	-0.43
12	0.3222	-0.8235	-0.1111	0.1044	-0.6429	-0.6089	-0.6194	-0.7373	-0.6239	-0.5556	-0.3075	-0.7827	-0.7908	-0.8667	-0.68
13	0.3264	-0.9176	-0.3333	-0.4710	-0.9244	-1	0.2388	-0.9322	-0.6506	0.1111	-0.3971	-0.9936	0.3936	-0.9394	-0.67
14	0.1281	-1	0.7778	-0.0534	-0.5840	-0.8711	0.5571	-0.7797	0.6866	-0.1111	-0.0591	-0.9617	0.7376	-0.9697	0.96

Figure 6: Normalized data.

After cluster analysis was performed using the data stream recovery and filtering method under conditions of intersecting clusters, recommendations are provided to improve the characteristics and eliminate the negative impact of its use, depending on the class, in accordance with drinking water standards. The method's quality was assessed using clustering quality indices. Table 3 presents the quality indicators for data clustering using different methods.

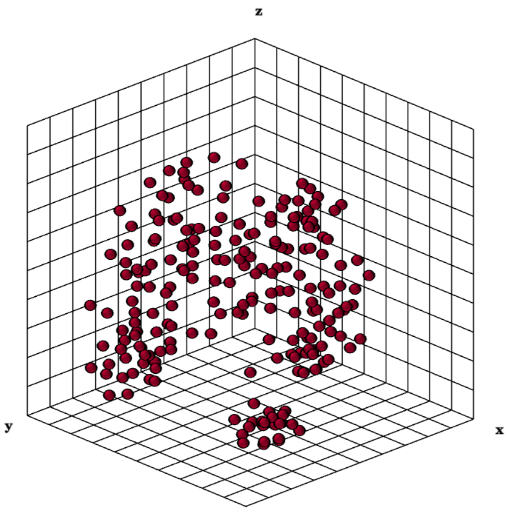


Figure 7: Scatter of observations in the space of a hypercube.

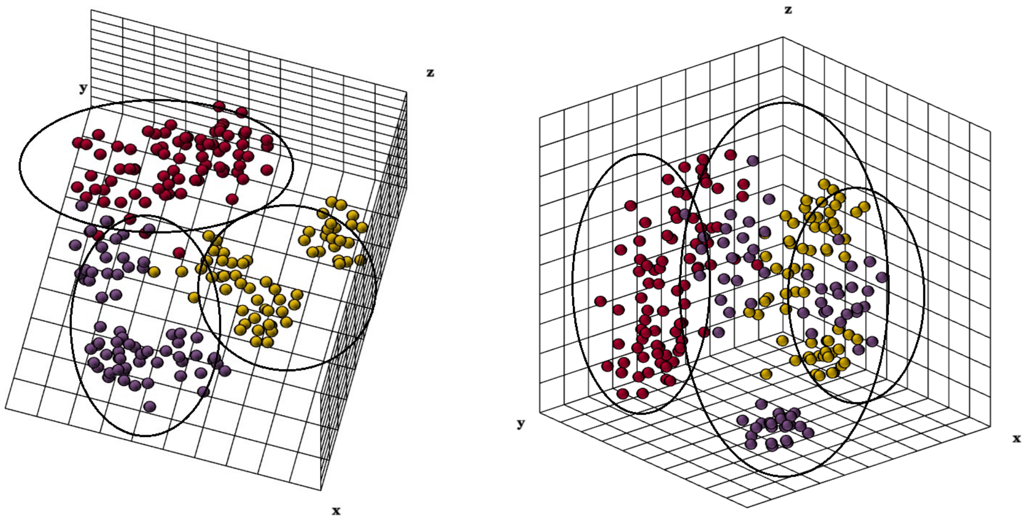


Figure 8: Dividing a sample of data into 3 clusters (demonstration from different angles).

Table 3
Assessing the quality of data clustering

Methods	PC	CE	SC	S	XB	DI
Neuro-fuzzy system ensembles based on credibilistic fuzzy clustering algorithms	0.87	0.46	7.60	2.72	5.61	0.27
FCM	0.79	0.55	7.33	0.54	5.51	0.11
Gustafson-Kessel	0.55	0.63	7.34	0.14	1.15	0.10

The proposed neuro-fuzzy ensembles provide a synergistic effect, combining the approximating capabilities of neural networks with the transparency of fuzzy logic. This allows us to recommend it as the most reliable tool for intelligent data analysis in decision support systems, where high classification accuracy with minimal entropy of results is critically important.

Conclusions

The problem of fuzzy clustering of data streams is considered under conditions when information is received online, possibly in real time, and the number of clusters is a priori unknown. It is proposed to use an ensemble of subsystems operating in parallel, each solving the fuzzy clustering problem online and designed for a different number of possible clusters. In this case, each subsystem solves the fuzzy trust clustering problem, and the quality of the results obtained is assessed using the fuzzy partition coefficient. It is assumed that the number of clusters during data processing can change arbitrarily. The proposed approach is computationally simple and intended for processing data that arrives at high frequency and can change its properties.

Acknowledgements

The work is supported by the state budget scientific research project of Kharkiv National University of Radio Electronics "Adaptive bagging of hybrid computational intelligence systems based on speed-optimal online learning" (SR No. 0124U000363).

Declaration on Generative AI

The author has not employed any Generative AI tools.

References

[1] Höppner, Frank, et al. Fuzzy cluster analysis: methods for classification, data analysis and image recognition. John Wiley & Sons, 1999.

[2] Zhang, Yong, et al. "Possibilistic c-means clustering based on the nearest-neighbour isolation similarity." Journal of Intelligent & Fuzzy Systems 44.2 (2023): 1781-1792.

[3] Gan, G., Ma, C., & Wu, J. (2020). Data clustering: theory, algorithms, and applications. Society for Industrial and Applied Mathematics.

[4] Raghavan, V., & Hafez, A. (2000, June). Dynamic data mining. In International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (pp. 220-229). Berlin, Heidelberg: Springer Berlin Heidelberg.

[5] Gaber, M. M. (2012). Advances in data stream mining. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2(1), 79-85.

[6] Gaber, M. M., Zaslavsky, A., & Krishnaswamy, S. (2010). Data stream mining. Data mining and knowledge discovery handbook, 759-787.

[7] Bodyanskiy, Y. V., Shafronenko, A., & Rudenko, D. (2019). Online neuro fuzzy clustering of data with omissions and outliers based on completion strategy. CEUR-WS, (pp. 18-27)

[8] Shafronenko, A. Y., Kasatkina, N. V., Bodyanskiy, Y. V., & Shafronenko, Y. O. (2023). Credibilistic robust online fuzzy clustering in data stream mining tasks. Radio Electronics, Computer Science, Control, (3), 97-103. DOI: 10.15588/1607-3274-2021-1-10

[9] T. Kohonen, Self-Organizing Maps. Berlin: Springer, 1995, 362 p. DOI: 10.1007/978-3-642-56927-2.

[10] Wongkhuenkaew, Ritipong, et al. "Fuzzy K-nearest neighbor based dental fluorosis classification using multi-prototype unsupervised possibilistic fuzzy clustering via cuckoo search algorithm." International Journal of Environmental Research and Public Health 20.4 (2023): 3394.

[11] Zhang, Yong, et al. "Possibilistic c-means clustering based on the nearest-neighbour isolation similarity." Journal of Intelligent & Fuzzy Systems 44.2 (2023): 1781-1792.

[12] Shafronenko A., Bodyanskiy Ye., Klymova I., Holovin O., Online credibilistic fuzzy clustering of data using membership functions of special type. Proceedings of The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020), April 27-1 May 2020. Zaporizhzhia, 2020. Access mode: http://ceur-ws.org/Vol-2608/paper56.pdf.

[13] J. Zhou, Credibilistic clustering: the model and algorithms. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. 2015. Vol. 23, №4. pp. 545-564. DOI: 10.1142/S0218488515500245.

[14] J. Zhou, Credibilistic clustering algorithms via alternating cluster estimation. Journal of Intelligent Manufacturing. 2017. Vol. 28. pp.727-738. DOI: 10.1007/s10845-014-1004-6.

[15] Shafronenko A., Bodyanskiy Ye., Klymova I., Holovin O., Online credibilistic fuzzy clustering of data using membership functions of special type. Proceedings of The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020), April 27-1 May 2020. Zaporizhzhia, 2020. Access mode: http://ceur-ws.org/Vol-2608/paper56.pdf.

[16] Bezdek, James C., Robert Ehrlich, and William Full. "FCM: The fuzzy c-means clustering algorithm." Computers & geosciences 10.2-3 (1984): 191-203.

[17] Bodyanskiy, Y., Vynokurova, O., Pliss, I., Setlak, G., & Mulesa, P. (2016, August). Fast learning algorithm for deep evolving GMDH-SVM neural network in data stream mining tasks. In 2016 IEEE First International Conference on Data Stream Mining & Processing (DSMP) (pp. 257-262). IEEE.

[18] Cardoso, Margarida GMS, and André Ponce de Leon F. De Carvalho. "Quality indices for (practical) clustering evaluation." Intelligent Data Analysis 13.5 (2009): 725-740.